

# **ESTUDIO DE LOS CONDICIONANTES QUE AFECTAN AL APRENDIZAJE DE LAS MATEMÁTICAS EMPRESARIALES.**

**Autores: Camacho Peñalosa, M.E.; Fernández Geniz, P.; Gómez Domínguez, D.;  
Masero Moreno, I.; Vázquez Cueto, M.J. y Zapata Reina, A.**

Facultad de Ciencias Económicas y Empresariales  
Departamento de Economía Aplicada III  
Avda. Ramón y Cajal, nº 1  
41018- SEVILLA  
Tfno: 954557553-954551673- 954557556  
Fax: 954551667  
e-mail: [pepi@us.es](mailto:pepi@us.es),  
[imasero@us.es](mailto:imasero@us.es).

## **Resumen**

El propósito de este trabajo es tratar de identificar la relevancia de algunas variables, características personales y de formación del alumno, sobre la calificación obtenida por el alumno en la asignatura Matemáticas I (programa de Álgebra Lineal), para la determinación, a priori, de la mayor o menor probabilidad de superar las asignaturas de Matemáticas de primer curso de la Licenciatura de Administración y Dirección de Empresas. Dado que la variable dependiente es discreta, se utilizarán las técnicas de regresión logística.

Palabras claves: Regresión Logística, Matemáticas Empresariales, Significación de variables, Análisis Discriminante, Clasificación.

## **Abstract**

In this paper, we try to measure the effect of some variables, personal data and previous learning of students, in the probability to pass the first level of mathematics for Administration and Business. We use two methods to discriminate the dependent variable, the logit analysis and the discriminant analysis.

Keywords: Regression, Logit Analysis, Mathematics, Discriminant Analysis, Classification.

## 1.- INTRODUCCIÓN.

Dos de los problemas que atañen a la enseñanza en la Universidad española actualmente son la heterogeneidad del nivel de conocimientos alcanzado por los alumnos en los estudios previos y las carencias de los mismos que se aprecian en materias concretas. Esto nos lleva a indagar en los posibles condicionantes de las deficiencias en conocimientos matemáticos observadas en los alumnos que se incorporan en los últimos años a los estudios universitarios de tipo económico-empresarial<sup>1</sup>. En líneas generales, esto puede llevar al alumno a obtener peores resultados en las calificaciones de las pruebas de evaluación<sup>2</sup> y a que se produzca un aumento de los no presentados a dichas pruebas.

Entre los aspectos que pueden influir en calificaciones de las asignaturas de matemáticas de la Licenciatura de Administración y Dirección de Empresas<sup>3</sup>, y en su decisión de abandonar las mismas, sobre todo el primer año en el que se cursan, se encuentra el bachillerato de procedencia, ya que, condiciona los conocimientos previos del alumno y, por tanto, el aprendizaje de las asignaturas de matemáticas .

No menos importantes son ciertos cambios que afectan al estudiante cuando se incorpora a la vida universitaria, por ejemplo, existe un menor control por parte de la institución educativa<sup>4</sup>. Además, hay que contemplar que al iniciar sus estudios universitarios, el alumno ha de cambiar y adaptar su forma de estudiar Matemáticas a la nueva situación, ya que se produce un cambio en su mecánica habitual de recepción de conocimientos, lo que añade mayor dificultad al estudio.

En algunos casos, se produce un menor control por parte del entorno familiar, que suele ser aún mayor cuando el alumno abandona la residencia familiar para iniciar sus estudios universitarios.

Algunos alumnos recurren a algún tipo de apoyo extra para la preparación de las asignaturas, ya sea acudiendo a algún centro no universitario que las imparte, “academia”, o acudiendo a un profesor particular.

Estos aspectos se han tenido en cuenta a la hora de realizar el siguiente estudio, en el que tratamos de analizar si influyen en los resultados obtenidos por los alumnos en las disciplinas de matemáticas. El trabajo se estructura en los siguientes apartados. En los apartados dos y tres se aborda la descripción de las variables consideradas y la selección de la muestra. En los dos apartados siguientes se aplican las técnicas de clasificación: en el apartado cuatro se realiza un análisis discriminante y en el apartado cinco, un análisis logit. En el último apartado se recogen las reflexiones finales.

## 2.- VARIABLES Y FUENTE DE DATOS.

Teniendo en cuenta las consideraciones recogidas en el apartado anterior definimos cinco variables.

**Variable Bachillerato Cursado.** Se han considerado cinco modalidades a las que se les han asignado los siguientes valores: Bachillerato de Arte (1), Bachillerato en Ciencias

de la Naturaleza (2), Bachillerato en Humanidades y Ciencias Sociales (3), Bachillerato en Tecnología (4) y otros (5).

Aunque lo más común es que nuestros alumnos procedan de la opción de Bachillerato en Ciencias Sociales y Humanidades, hay alumnos que proceden de otras opciones. Además, debemos tener en cuenta que no sólo va a influir el bachillerato cursado sino también el itinerario curricular elegido por el alumno, ya que sus conocimientos dependerán de las asignaturas de matemáticas que haya cursado. Introducimos, por tanto, una segunda variable que tendrá en cuenta este condicionante.

**Variable Asignaturas de matemáticas cursadas en los dos años de bachillerato** asignando los siguientes valores: Ninguna (0), sólo han cursado la de primer curso (1), sólo han cursado la de segundo curso (2), han cursado matemáticas los dos años de bachillerato (3).

Hemos de considerar que algunos de los cambios que se producen en el alumno o su entorno cuando inicia los estudios universitarios, no son fáciles de recoger en un estudio de este tipo. Nosotros sólo vamos a incidir en la mayor autonomía o “menor control” de los padres de aquellos alumnos que abandonan la residencia familiar al iniciar sus estudios universitarios. Tenemos en cuenta, por tanto, una tercera variable.

**Variable Vive con la Familia:** El alumno vive en el domicilio familiar (1), no vive en el domicilio familiar (0).

La cuarta variable considera si el alumno dispone o no de apoyo extra en la preparación de las asignaturas de matemáticas.

**Variable Apoyo Extra:** El alumno recurre a una “academia” o profesor particular para la preparación de las matemáticas (1), no cuenta con este apoyo (2).

Tomamos una quinta variable, para analizar si el sexo puede condicionar los resultados del alumno.

**Variable Sexo:** Mujer (0) y hombre (1)

La variable dependiente que tratamos de explicar es haber superado o no la asignatura de Matemáticas I<sup>5</sup>, variable a la que se le ha asignado los valores 1 y 0 respectivamente.

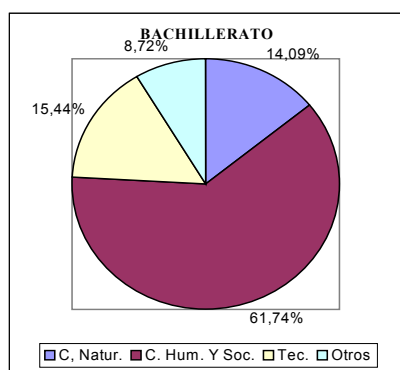
Para la obtención de los datos, al inicio del segundo semestre del presente curso (2002/03), se les facilitó a los alumnos de la asignatura Matemáticas II de la Licenciatura de Administración y Dirección de Empresas una ficha en la que se les solicitaba información sobre sus datos personales, el bachillerato elegido, asignaturas de matemáticas cursadas en el mismo, apoyo en la preparación de las asignaturas de matemáticas de la licenciatura, acceso a internet en el domicilio, calificación obtenida en la asignatura Matemáticas I y además un breve cuestionario para evaluar sus conocimientos previos.

Tras la recogida de las fichas se procedió a la creación de una base de datos sobre la que se extrae la muestra que se utilizará en este estudio.

En total, en la asignatura de Matemáticas II hay matriculados 751 alumnos y la muestra consta de 149 de estos alumnos, de los cuales se dispone de información sobre todas las variables seleccionadas.

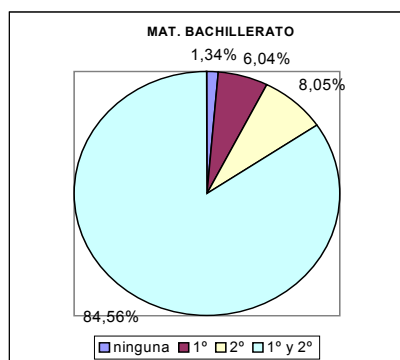
### 3.- ESTUDIO DE LA MUESTRA.

De los datos recogidos en las fichas entregadas por los alumnos de nuestra muestra, vamos a presentar las siguientes tablas en las que se recoge la distribución de los alumnos en porcentajes para cada una de las variables de este estudio.



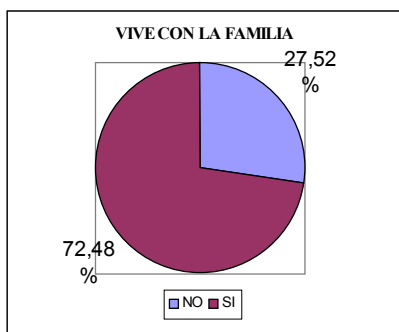
VARIABLE BACHILLERATO CURSADO		
CATEGORÍAS	FRECUENCIA	%
B. Arte	0	0
B. Cc. de la Naturaleza	21	14,09
B. Hum. y CC. Sociales	92	61,74
B. Tecnología	23	15,44
Otros	13	8,72

En cuanto al bachillerato cursado por los alumnos de la muestra, un 61,74% proceden del Bachillerato de ciencias Sociales y Humanidades, seguido a gran distancia por el de Tecnología (15,44%) y el de Ciencias de la Naturaleza (14,09%). Hemos de destacar, que curiosamente, la preparación matemática proporcionada por estos dos bachilleratos es muy superior a la que proporciona el de Ciencias Sociales y Humanidades, que es el más numeroso en nuestra muestra. Además, hay que destacar el hecho de que hay un 8,72% de la muestra que procede de los antiguos estudios de COU, es decir, que aún quedan alumnos en primer curso que cursaron estudios medios extinguidos en el curso 2001-2002.



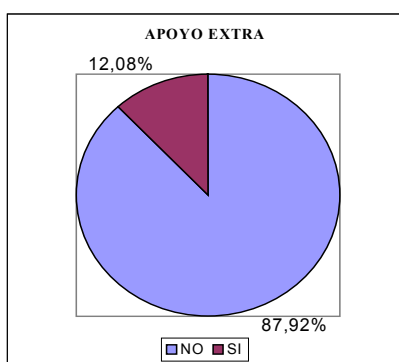
VARIABLE MATEMÁTICAS CURSADAS EN BACHILLERATO		
CATEGORÍAS	FRECUENCIA	%
NINGUNA	2	1,34
EN PRIMERO	3	6,04
EN SEGUNDO	12	8,05
EN AMBOS	126	84,56

La gran mayoría de los alumnos de la muestra ha cursado matemáticas durante los dos años de bachillerato.



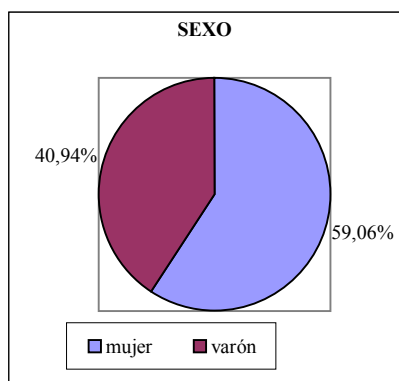
VARIABLE VIVE CON LA FAMILIA		
CATEGORÍAS	FRECUENCIA	%
SI	108	72,48
NO	41	27,52

El 72,48% de los alumnos de la muestra conviven con la familia, frente a un 27,52% cuyo domicilio durante el curso se encuentra en Sevilla y no es el familiar.



VARIABLE APOYO EXTRA		
CATEGORÍAS	FRECUENCIA	%
SI	131	87,92
NO	18	12,08

Destaca que sólo el 12,08% del alumnado de nuestra muestra utiliza apoyo extra para las asignaturas de Matemáticas.



VARIABLE SEXO		
CATEGORÍAS	FRECUENCIA	%
VARON	61	40,94
MUJER	88	59,06

Como puede observarse, en la muestra el porcentaje de mujeres es superior al de los varones.

El alumno tipo de la muestra, pues, ha cursado Bachillerato de Ciencias Sociales con las Matemáticas de primer y segundo curso, vive con su familia, no utiliza apoyo extra y es mujer.

Tras analizar cada variable de forma independiente, hemos estudiado el nivel de correlación existente entre las variables consideradas. Para ello, se ha aplicado el test de correlación de Kendall a pares de variables, obteniéndose como resultado que, únicamente son significativas, a un nivel de significación superior al 99'9%, las correlaciones existentes entre las siguientes variables:

- Apoyo extra y calificación en Matemáticas I.
- Apoyo extra y bachillerato cursado.
- Apoyo extra y número de asignaturas de matemáticas cursadas en Bachillerato.
- Apoyo extra y sexo.

A continuación vamos a tratar de clasificar, a priori, a los alumnos en dos grupos: los que han superado la asignatura de Matemáticas I y los que no la han superado, en función de los valores que toman las cinco variables consideradas en el estudio.

#### 4.- ANÁLISIS DISCRIMINANTE.

El análisis discriminante es una técnica estadística que permite asignar una observación a una de varias poblaciones<sup>6</sup>. Parte de la existencia de unos grupos constituidos cada uno por un conjunto de individuos. Para el caso que nos ocupa, ya sea el estudio del éxito o el fracaso académico, dispondríamos de dos grupos: alumnos aprobados y no aprobados, incluyendo en esta categoría a los no presentados. Estos individuos vienen caracterizados por un conjunto de variables  $x_1, x_2, \dots, x_n$ , así el individuo  $X_i$  viene representado por el siguiente vector  $(x_{i1}, x_{i2}, \dots, x_{in})$ .

Este tipo de análisis puede utilizarse con una finalidad puramente descriptiva, es decir tratando de explicar la pertenencia de un individuo a un grupo, comprobando si el modelo lo asigna o no correctamente y tratando de cuantificar el peso que cada variable tiene en la discriminación. Ahora bien, este tipo de análisis también tiene una aplicación con el objeto de hacer predicciones sobre la pertenencia de un nuevo individuo con un perfil de variables a uno de los grupos.

Supongamos que el éxito viene determinada por un conjunto de variables:

$$Z = f(x_1, x_2, \dots, x_n)$$

donde  $x_i$  con  $i = 1, 2 \dots n$ , son las variables independientes y  $Z$  es la variable dependiente que podría interpretarse como una variable global, síntesis de las utilizadas. A la expresión anterior se denomina función discriminante.

Dada una observación perteneciente a la muestra de alumnos necesitamos disponer de una regla de clasificación apropiada, de forma que para cada alumno caracterizado por su vector de observaciones  $X_i = (x_{i1}, x_{i2}, \dots, x_{in})^t$ , se calcula el valor  $Z_i$ , si  $Z_i < Z_c$  esta función clasifica al alumno como aprobado.  $Z_c$  es la puntuación crítica o puntuación de corte. Una forma de determinarla es calculando la media aritmética de los valores que toma la función en los puntos formados por las medias de las variables en ambos grupos, es decir:

$$Z_c = \frac{Z_A + Z_{NA}}{2}$$

siendo:

$Z_A$ : el valor de la función para la media de las variables en el grupo de aprobados.

$Z_{NA}$ : el valor de la función para la media de las variables en el grupo de no aprobados.

La función discriminante lineal adopta la forma:

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (2)$$

Existen distintos métodos para la obtención de la misma. La forma más comúnmente utilizada para estimar los parámetros de la función discriminante ha sido el método de máxima verosimilitud.

En la aplicación de la regla discriminante se pueden cometer dos tipos de errores, los denominados error Tipo I (asignar un alumno del grupo de aprobados al grupo de no aprobados) y error Tipo II (asignar un alumno al grupo de aprobados cuando en realidad es del grupo de no aprobados).

Para acometer nuestro análisis, utilizamos el paquete estadístico XLSTAT, versión 6.0.

Hemos realizado el estudio con las cinco variables consideradas en el trabajo: Sexo (V1), convivencia en el domicilio familiar (V2), modalidad de bachillerato (V3), apoyo académico extra (V4), y asignaturas de matemáticas cursadas en el bachillerato (V5).

Mediante el test de Box, que se distribuye según una F-Fisher, rechazamos la hipótesis nula de igualdad de las matrices de varianza-covarianza en los dos grupos, y, por tanto, lo apropiado sería el análisis discriminante cuadrático.

Teniendo en cuenta la consideración anterior, la salida del programa, recogida en la siguiente tabla, son dos funciones discriminantes:

	Alumnos No Aprobados	Alumnos Aprobados
Coeficiente	0	1
Constante	-19.66	-22.971
Sexo	0.659	1.172
Vive con la familia	5.407	6.232
Bachillerato	5.267	5.666
Apoyo extra	1.353	2.42
Matemáticas cursadas	7.37	7.803

Para cada observación se calcula la puntuación con las dos funciones, y se asigna la observación a la de mayor valor, resultando la siguiente tabla de clasificación.

	Clasificados como no aprobados	Clasificados como aprobados	Total
Alumnos no aprobados	24 16,11%	21 14,09%	45 30,20%
Alumnos aprobados	28 18,79%	76 51,01%	104 69,80%
Total	52 34,90%	97 65,10%	149 100%

Tasa de error global: 32,89%

## 5.- ANÁLISIS LOGIT

Este tipo de modelo, también llamado modelo de respuesta cualitativa, tiene utilidad para pronosticar que sucederá cuando existen dos posibilidades, en nuestro caso, superar la asignatura o no superarla.

Supongamos que la superación de la asignatura por parte del alumno está asociada a un índice de éxito ( $G^*$ ), de manera que cuando supera un valor determinado,  $G^* > G^0$ , se produce el éxito, siendo  $G^0$  el valor o punto crítico.

Pese a que el índice de éxito no es directamente observable, a priori, vamos a suponer que depende linealmente de un conjunto de indicadores ( $x_1, x_2, \dots, x_k$ ) que sí son directamente observables y de una variable aleatoria, denotada por  $u$ . Modelizamos el índice de éxito como:

$$G^* = \sum_{i=1}^k \beta_i x_i + u = X^t \beta + u$$

Si se define  $G = G^* - G^0$ , entonces el éxito se produce cuando  $G > 0$ . El modelo transformado es:

$$G = X^t \beta + u - G^0 = \dot{X}^t \dot{\beta} + u$$

donde:

$$\dot{X}^t = (1 \ x_1 \ x_2 \ \dots \ x_k); \quad \dot{\beta} = (\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k)^t \quad \text{y} \quad \beta_0 = -G^0$$

Partiendo del índice de éxito establecido, la probabilidad de éxito viene dada por:

$$P(\text{éxito}) = P(G > 0) = P(\dot{X}^t \dot{\beta} + u > 0) = 1 - P(u \leq -\dot{X}^t \dot{\beta})$$

por tanto, la probabilidad de que un alumno supere la asignatura depende de que una variable aleatoria ( $u$ ) no sobrepase un determinado valor. Es decir, depende del valor en un punto de la función de distribución de una variable aleatoria.

La probabilidad de no éxito sería:

$$P(\text{no éxito}) = 1 - P(\text{éxito}) = P(u \leq -\dot{X}^t \dot{\beta}) = F(-\dot{X}^t \dot{\beta})$$

El problema ahora consiste en plantear un modelo apropiado para el último miembro de la ecuación. Existen tres posibilidades que coinciden con los tres modelos de elección discreta más comunes: modelo lineal de probabilidad, modelo probit y modelo logit<sup>7</sup>.

Si la función de distribución elegida para la variable aleatoria " $u$ " es la función de distribución de la ley de probabilidad logística estándar, el modelo especificado es el conocido como modelo logístico o logit, es decir:



$$P(Z=1) = \frac{1}{1 + \exp(\dot{X}^t \dot{\beta})} \quad P(Z=0) = \frac{1}{1 + \exp(-\dot{X}^t \dot{\beta})}$$

donde Z es la variable dependiente.

Para la estimación de los coeficientes  $\beta_i$  utilizamos el método de máxima verosimilitud.

El análisis logit presenta la ventaja de no requerir que las variables se distribuyan según la ley normal, lo que muchos argumentan como la razón fundamental para que este enfoque resulte mucho más robusto que el enfoque discriminante.

Con el fin de conseguir una función que permita predecir la probabilidad de aprobar la asignatura de Matemáticas I, en función del valor de las cinco variables consideradas, hemos aplicado a los 149 datos de la muestra el análisis logit. La función de probabilidad obtenida es:

$$P(z=1) = \frac{1}{1 + \exp(-2,401^* + 0,519V1 + 0,789^* V2 + 0,397V3 + 1,346^* V4 + 0,409V5)}$$

\* con un nivel de significación mayor del 5%.

El paquete utilizado proporciona también una medida de la bondad del ajuste. En este caso,  $R^2$  (McFadden)= 0,076.

La tabla de clasificación análoga a la presentada para el análisis discriminante arroja los siguientes resultados:

	Clasificados como no aprobados	Clasificados como aprobados	Total
Alumnos no aprobados	6 4,03%	39 26,17%	45 30,20%
Alumnos aprobados	4 2,69%	100 67,11%	104 69,80%
Total	10 7,72%	139 92,28%	149 100%

Aunque el ajuste no es bueno, el programa XLSTAT proporciona un estadístico que mide el poder de información significativa de las variables incluidas en el modelo. En este caso, el estadístico, que se distribuye según una  $\chi^2$  con 5 grados de libertad, tiene un valor de 13,901, lo que implica una probabilidad de 0,016 de estar equivocado cuando decimos que las variables explicativas dan información significativa.

## 6.- REFLEXIONES FINALES.

De todo lo expuesto en los apartados anteriores, podemos deducir que las técnicas de clasificación son apropiadas para determinar a priori la probabilidad de que un alumno

supere una determinada disciplina, en nuestro caso concreto, la asignatura de Matemáticas I de la Licenciatura de Administración y Dirección de Empresas.

No existe diferencia significativa entre el porcentaje de clasificación correcta obtenido cuando se aplica el análisis discriminante y el obtenido cuando se aplica el análisis logit.

Las variables consideradas en el estudio son significativas, pero los estadísticos que miden la bondad del ajuste de los modelos toman valores muy bajos, lo que nos informa de que no son éstos los apropiados.

En un estudio posterior, trataremos de corregir los resultados anteriores, introduciendo información complementaria acerca de nuestros alumnos.

#### Bibliografía:

- Anemia, T. (1981). Qualitative Response Models: A Survey. *Journal of Economic Literature*, Vol.19, num.4, pp. 1483-1535.
- Altman, E.I. et alia. (1981). Application of Classification Techniques in Business, Banking and Finance. *Contemporary Studies in Economic and Financial Analysis*, Vol. 3, Jai Press, Greenwich.
- Cuadras, C. M. (1991). Métodos del Análisis Multivariante. Ed. PPU, Barcelona.
- Fernández, P. et al. (2000). Algunas consideraciones sobre la influencia de los estudios previos en las calificaciones de Matemáticas para los alumnos de la Diplomatura de Empresariales. VIII Jornadas ASEPUMA, Sevilla, pp. 117-128.
- Fernández, P. Et al. (2001). Reflexiones sobre la adecuación de los conocimientos previos de los alumnos al temario de Matemáticas de la Escuela de Empresariales. IX Jornadas ASEPUMA, Las Palmas de Gran Canaria.
- Greene, W.H. (1999). Análisis Econométrico. Prentice-Hall, 3ª edición, Madrid. (traducción de la 3ª edición inglesa de 1998).
- Lachenbruch, P.A. (1975). Discriminant Análisis. Hafner Press, New York.
- Maddala, G.S. (1983). Limited-Dependent and Quantitative Variables in Econometrics. Cambridge University Press, Cambridge.
- Novales Cinca, A. (1993). Econometría. McGraw-Hill (2ª Edición), Madrid.

---

<sup>1</sup> En esta línea hemos presentado otros trabajos, Fernández et al. 2000 y 2001.

<sup>2</sup> Que suele ser por lo general el examen tradicional.

<sup>3</sup> De igual forma se podría conducir el estudio para la Licenciatura de Economía y para la Diplomatura de Ciencias Empresariales.

<sup>4</sup> Por ejemplo, no es habitual el control de la asistencia.

<sup>5</sup> Esta asignatura se cursa en el primer semestre de la Licenciatura de Administración y Dirección de Empresas y sus contenidos son básicamente de Álgebra Lineal.

<sup>6</sup> Un análisis detallado de esta técnica puede encontrarse en Cuadras(1991), Lachenbruch (1975) y Altman et Alia (1981).

<sup>7</sup> Para mayor detalle sobre los modelos de elección discreta o respuesta cualitativa, pueden consultarse, por ejemplo, las siguientes publicaciones: Novales (1993, fundamentalmente el cap. 16); Maddala (1983, cap. 2 y 5); Anemiya (1981), y Greene (1997, cap. 19).